DTIC FILE COPY

(2)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

AD-A218 566

1a. REPORT SECURITY CLASSIFICATION
Unclassified

2a. SECURITY CLASSIFICATION AUTHORITY
N/A

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE
N/A

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for Public Release;
Distribution Unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
Technical Report No. 279

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 90-0274

6a. NAME OF PERFORMING ORGANIZATION
University of North Carolina
Center for Stochastic Processes

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
AFOSR/NM

6c. ADDRESS (City, State, and ZIP Code)
Statistics Department
CB #3260, Phillips Hall
Chapel Hill, NC 27599-3260

7b. ADDRESS (City, State, and ZIP Code)
Bldg. 410
Bolling Air Force Base, DC 20332-6448

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
AFOSR

8b. OFFICE SYMBOL (If applicable)
NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
F49620 85C 0144

8c. ADDRESS (City, State, and ZIP Code)
Bldg. 410
Bolling AFB, DC 20332-6448

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.
6.1102F	2304	A5	

11. TITLE (Include Security Classification)
Parameter estimation in linear filtering

12. PERSONAL AUTHOR(S)
Kallianpur, G. and Selukar, R.S.

13a. TYPE OF REPORT
preprint

13b. TIME COVERED
FROM ________ TO ______

14. DATE OF REPORT (Year, Month, Day)
October 1989

15. PAGE COUNT
31

16. SUPPLEMENTARY NOTATION
None .

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
XXXXXX	XXXXXXXXXX	XXX

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Kalman filtering, identification, large deviation

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

See over.

DTIC
ELECTE
FEB 27 1990
S D
D

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
Professor Eytan Barouch

22b. TELEPHONE (Include Area Code)
(202) 767-~~5026~~ 4940

22c. OFFICE SYMBOL
AFOSR/NM

DD Form 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFED

UNCLASSIFIED

19. Abstract.

Suppose on a probability space $(\Omega, \mathbf{F}, \mathbf{P})$ a partially observable random process $(x_t, y_t), t \geq 0$; is given where only the second component (y_t) is observed. Furthermore assume that (x_t, y_t) satisfy the following system of stochastic differential equations driven by independent Wiener processes $(W_1(t))$ and $(W_2(t))$:

$$dx_t = -\beta x_t \, dt + dW_1(t), \qquad x_0 = 0$$

$$dy_t = \alpha x_t \, dt + dW_2(t), \qquad y_0 = 0; \quad \alpha, \beta \in (a, b), \; a > 0.$$

We obtain a large deviation inequality for the maximum likelihood estimator (m.l.e.) of the unknown parameter $\theta = (\alpha, \beta)$. This inequality enables us to prove the strong consistency, asymptotic normality and covergence of the moments of the m.l.e.. The method of proof can be extended to obtain similar results when multi-dimensional instead of one dimensional processes are considered and θ is a k-dimensional vector.

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

PARAMETER ESTIMATION IN LINEAR FILTERING

by

G. Kallianpur and R.S. Selukar

Technical Report No. 279

October 1989

219. J. Bather. Stopping rules and ordered families of distributions. Dec. 87 *Sequential Anal.*, 7, 1988, 111-126.

220. S. Cambanis and M. Maejima. Two classes of self-similar stable processes with stationary increments. Jan. 88. *Stochastic Proc. Appl.*, 32, 1989, 305-329.

221. H.P. Hucke, G. Kallianpur and R.L. Karandikar. Smoothness properties of the conditional expectation in finitely additive white noise filtering. Jan. 88. *J. Multivariate Anal.*, 27, 1988, 261-269.

222. I. Mitoma. Weak solution of the Langevin equation on a generalized functional space. Feb. 88. (Revised as No. 238).

223. L. de Haan, S.I. Resnick, H. Rootzén and C. de Vries. Extremal behaviour of solutions to a stochastic difference equation with applications to arch-processes. Feb. 88. *Stochastic Proc. Appl.*, 32, 1989, 213-224.

224. O. Kallenberg and J. Szulga. Multiple integration with respect to Poisson and Lévy processes. Feb. 88. *Prob. Theor. Rel. Fields*, 83, 1989, 101-134.

225. D.A. Dawson and L.G. Gorostiza. Generalized solutions of a class of nuclear space valued stochastic evolution equations. Feb. 88 *Appl. Math. Optimization*, to appear.

226. G. Samorodnitsky and J. Szulga. An asymptotic evaluation of the tail of a multiple symmetric α-stable integral. Feb. 88. *Ann. Probability*, 17, 1989, 1503-1520.

227. J.J. Hunter. The computation of stationary distributions of Markov chains through perturbations. Mar. 88.

228. H.C. Ho and T.C. Sun. Limiting distribution of nonlinear vector functions of stationary Gaussian processes. Mar. 88. *Ann. Probability*, to appear.

229. R. Brigola. On functional estimates for ill-posed linear problems. Apr. 88.

230. M.R. Leadbetter and S. Nandagopalan. On exceedance point processes for stationary sequences under mild oscillation restrictions. Apr. 88. *Proc. Oberwolfach Conf. on Extremal Value Theory*, J. Hüsler and R. Reiss, eds., Springer, 1989, 69-80.

231. S. Cambanis, J. P. Nolan and J. Rosinski. On the oscillation of infinitely divisible processes. Apr. 88. *Stochastic Proc. Appl.*, to appear.

232. G. Hardy, G. Kallianpur and S. Ramasubramanian. A nuclear space-valued stochastic differential equation driven by Poisson random measures. Apr. 88.

233. D.J. Daley, T. Rolski. Light traffic approximations in queues (II). May 88. *Math. Operat. Res.*, to appear.

234. G. Kallianpur, I. Mitoma, R.L. Wolpert. Nuclear space-valued diffusion equations. July 88. *Stochastics*, 1989, to appear.

235. S. Cambanis. Admissible translates of stable processes: A survey and some new models. July 88.

236. E. Platen. On a wide range exclusion process in random medium ... local jump intensity. Aug. 88.

237. R.L. Smith. A counterexample concerning the extremal index. Aug. 88. *Adv. Appl. Probab.*, 20, 1988 681-683.

238. G. Kallianpur and I. Mitoma. A Langevin-type stochastic differential equation on a space of generalized functionals. Aug. 88.

239. C. Houdré. Harmonizability, V-boundedness, (2,P)-boundedness of stochastic processes. Aug. 88. *Prob. Th. Rel. Fields*, to appear.

240. G.W. Johnson and G. Kallianpur. Some remarks on Hu and Meyer's paper and infinite dimensional calculus on finitely additive canonical Hilbert space. Sept. 88. *Th. Prob. Appl.*, to appear.

241. L. de Haan. A Brownian bridge connected with extreme values. Sept. 88. *Sankhya*, 1989, to appear.

242. O. Kallenberg. Exchangeable random measures in the plane. Sept. 88. *J. Theor. Probab.*, to appear.

243. E. Masry and S. Cambanis. Trapezoidal Monte Carlo integration. Sept. 88. *SIAM J. Numer. Anal.*, 1989, to appear.

244. L. Pitt. On a problem of H.P. McKean. Sept. 88. *Ann. Probability*, 17, 1989, 1651-1657.

245. C. Houdré. On the linear prediction of multivariate (2,P)-bounded processes. Sept. 88.

246. C. Houdré. Stochastic processes as Fourier integrals and dilation of vector measures. Sept. 88. *Bull. Amer. Math. Soc.*, to appear.

247. J. Mijnheer. On the rate of convergence in Strassen's functional law of the iterated logarithm. Sept. 88. *Probab. Theor. Rel. Fields*, to appear.

248. G. Kallianpur and V. Perez-Abreu. Weak convergence of solutions of stochastic evolution equations on nuclear spaces. Oct. 88. Proc. Trento Conf. on *Infinite Dimensional Stochastic Differential Equations*, 1989, to appear.

249. R.L. Smith. Bias and variance approximations for estimators of extreme quantiles. Nov. 88.

250. H. Hurd. Spectral coherence of nonstationary and transient stochastic processes. Nov. 88. 4th Annual ASSP Workshop on Spectrum Estimation and Modeling, Minneapolis, 1988, 387-390.

251. J. Leskow. Maximum likelihood estimator for almost periodic stochastic processes models. Dec. 88.

252. M.R. Leadbetter and T. Hsing. Limit theorems for strongly mixing stationary random measures. Jan. 89. *Stochastic Proc. Appl.*, to appear.

253. M.R. Leadbetter, I. Weissman, L. de Haan, H. Rootzén. On clustering of high values in statistically stationary series. Jan. 89. *Proc. 4th Int. Meeting on Stat. Climatology*, to appear.

254. J. Leskow. Least squares estimation in almost periodic point processes models. Feb. 89.

255. N.N. Vakhania. Orthogonal random vectors and the Hurwitz-Radon-Eckmann theorem. Apr. 89.

Parameter Estimation in Linear Filtering

by

G. Kallianpur and R. S. Selukar *

University of North Carolina at Chapel Hill

Suppose on a probability space $(\Omega, \mathbf{F}, \mathbf{P})$ a partially observable random process $(x_t, y_t), t \geq 0$; is given where only the second component (y_t) is observed. Furthermore assume that (x_t, y_t) satisfy the following system of stochastic differential equations driven by independent Wiener processes $(W_1(t))$ and $(W_2(t))$:

$$dx_t = -\beta x_t \, dt + dW_1(t), \qquad x_0 = 0$$

$$dy_t = \alpha x_t \, dt + dW_2(t), \qquad y_0 = 0; \quad \alpha, \beta \in (a, b), \; a > 0.$$

We obtain a large deviation inequality for the maximum likelihood estimator (m.l.e.) of the unknown parameter $\theta = (\alpha, \beta)$. This inequality enables us to prove the strong consistency, asymptotic normality and covergence of the moments of the m.l.e.. The method of proof can be extended to obtain similar results when multidimensional instead of one dimensional processes are considered and θ is a k-dimensional vector.

Key words: Kalman filtering, identification, large deviation

AMS 1980 subject classification: Primary 62F10; Secondary 62M20

* Now at the University of Illinois at Urbana-Champaign

Research partially supported by the Air Force Office of Scientific Research Contract No. F49620 85C 0144.

Parameter Estimation in Linear Filtering

I. Introduction Suppose on a probability space $(\Omega, \mathbf{F}, \mathbf{P})$ a partially observable random process $(x_t, y_t), t \geq 0$; is given where only the second component (y_t) is observed (both the components could be vector valued). Furthermore assume that (x_t, y_t) satisfy the following system of stochastic differential equations (SDE):

$$dx_t = F\, x_t\, dt + G\, dW_1(t), \qquad x_0 = X_0$$

$$dy_t = H\, x_t\, dt + dW_2(t), \qquad y_0 = 0; \tag{1.1}$$

where $(W_1(t))$ and $(W_2(t))$ are *independent* standard Wiener processes and F, G, H are nonrandom matrix valued functions of appropriate order. The initial value X_0 is assumed to be a Gaussian random variable independent of both $(W_1(t))$ and $(W_2(t))$.

The estimation of unknown parameters in H, G and F, based on observations $(y_t, 0 \leq t \leq T)$ is known as "system identification". It appears that this problem of system identification was first considered by Balakrishnan (1973), who proved the weak local consistency of the maximum likelihood estimator (m.l.e.) under suitable regularity and identifiability assumptions. Later Bagchi and Borkar(1984) showed the strong global consistency of the m.l.e., for a slightly more general model. In their case the signal process could be an infinite dimensional process of the following kind:

$$x_t = \int_0^t S_{t-s}\, D\, dW_1(s)$$

where $S_t, t \geq 0$, is a strongly continuous semigroup with generator A on a separable Hilbert space $\mathbf{H}$, W_1 is a Brownian motion on a separable Hilbert space $\mathbf{K}$, and D is a bounded linear operator from $\mathbf{K}$ to $\mathbf{H}$. The observation process y_t,

however, is finite dimensional and satisfies the following SDE:

$$y_t = \int_0^t C\, x_s\, ds + W_2(t)$$

where C is a bounded linear operator from $\mathbf{H}$ to R^q and W_2 is R^q valued Brownian motion independent of W_1. The vector of unknown system parameters θ is assumed to be a point from a compact set in R^k.

Under suitable stability, controllability and differentiability assumptions they prove the strong consistency of the m.l.e. of θ. No discussion of the rate of convergence is provided.

It was Kutoyants(1984) who first considered the question of asymptotic normality of the m.l.e. in this setting. However, he only considered the following special case of the model in (1.1):

$$dx_t = -\beta x_t\, dt + dW_1(t), \qquad x_0 = 0$$

$$dy_t = \alpha x_t\, dt + dW_2(t), \qquad y_0 = 0; \quad \alpha, \beta \in (a, b),\ a > 0. \tag{1.2}$$

(All the processes involved in (1.2) are assumed to be one dimensional). In the above model, *when* β *is a known constant,* he obtained a large deviation inequality for the m.l.e. of α which in turn implies the strong consistency, asymptotic normality and the convergence of moments.

Here we extend this result to the m.l.e. of the bivariate parameter $\theta = (\alpha, \beta)$. It should be emphasized that Kutoyants's technique can not be applied to this bivariate estimation problem (not even for the univariate estimation of β when α is a known constant) since a special type of dependence of the filtered signal (which appears in the likelihood ratio) on the unknown parameter α is very essential for his approach. In the case of the above model the dependence of the filtered signal on the unknown parameter θ is not of this particular type and thus his technique is no longer applicable. This comment is briefly explained in a remark (Remark 3.2)

at the end of this article. On the other hand it will be clear that the method we have used can be applied without any major modification to the general model considered in (1.1) if, besides identifiability, the following two conditions are satisfied:

i) The parameter space Θ is an open, *bounded* subset of R^k.

ii) The eigenvalues of the matrix F lie in the *open* left-half of the complex plane.

(However, the computations become quite cumbersome.)

The main result along with the necessary notation is given in Section 2 and the proof which is based on Theorem 3.1 is givenin Section 3.

2. Notation and Statements of Results From now on, unless mentioned otherwise, the signal and observation processes $x_t, y_t, t \geq 0$; will refer to the solution of the SDE in (1.2). Also assume that the bivariate parameter $\theta = (\alpha, \beta)$ is an element of $\Theta = (a,b) \times (a,b)$, $a > 0$, $b < \infty$. The letter C (with or without a subscript) will denote a positive constant *independent of T (the time parameter)*; it need not be the same in two different expressions.

For $0 \leq t$ let $\hat{x}_t$ be the conditional expectation of x_t given the observations up to time t i.e.,

$$\hat{x}_t = E(x_t \mid \mathbf{F}^y{}_t) \tag{2.1}$$

where $\mathbf{F}^y{}_t$ is the σ- field generated by $\{ y_s, 0 \leq s \leq t \}$ and all the $\mathbf{P}$-null sets; furhermore let

$$d\nu_t = dy_t - \alpha \hat{x}_t \, dt. \tag{2.2}$$

Then it is well-known that (ν_t) is a Wiener process and moreover the process $(\hat{x}_t)$ satisfies the SDE,

$$d\hat{x}_t = -\beta \hat{x}_t \, dt + \alpha a_t \, d\nu_t, \qquad \hat{x}_0 = 0;$$

where a_t is the (unique) solution of a (deterministic) differential equation known as the Riccati equation. More precisely, a_t is the solution of the following

nonlinear differential equation:

$$\frac{d}{dt}a_t = 1 - \alpha^2 a_t^{\,2} - 2\beta\, a_t, \qquad a_0 = 0.$$

It is also known that as $t \to \infty$, $a_t \to a_\theta$, $a_\theta = \dfrac{-\beta + \sqrt{\alpha^2+\beta^2}}{\alpha^2}$. All these facts can be found in Liptser & Shiryayev (1978, Vol II, 16.2). Now we make a symplifying assumption commonly made in the literature (see, e.g., Kutoyants (1984, pp 103)). We assume that the system has reached the steady state i.e. we assume that $\hat{x}_t$ satisfies the SDE given below:

$$d\hat{x}_t = -\beta\,\hat{x}_t\, dt + \alpha\, a_\theta\, d\nu_t, \qquad \hat{x}_0 = 0. \tag{2.3}$$

Then from (2.2) and (2.3), it is easy to verify that

$$\hat{x}_t(\theta, y) = \alpha^2 a_\theta \int_0^t e^{-b_\theta(t-s)}\, dy_s \tag{2.4}$$

where $b_\theta = \sqrt{\alpha^2+\beta^2}$.

Let C_T denote the space of real valued continuous functions defined on $[0,T]$ endowed with the sup-norm topology and let $\mathbf{C}_T$ be the σ-field of Borel sets in C_T. Furthermore, let $P_T^{\,\theta}$ denote the measure induced by the paths $(y_s, 0 \le s \le T)$ on $(C_T, \mathbf{C}_T)$.

Then in view of the relation

$$dy_t = \alpha\,\hat{x}_t(\theta, y)\, dt + d\nu_t^{\,\theta}, \tag{2.5}$$

and the fact that $\nu_t^{\,\theta}$ is a Wiener process, it follows that $P_T^{\,\theta}$ is equivalent to the standard Wiener measure μ_W defined on $(C_T, \mathbf{C}_T)$. Furhermore the density or the likelihood function of the data $(y_s, 0 \le s \le T)$ at θ is given by

$$\frac{dP_\theta^{\,T}}{d\mu_W}(y) = \exp\left(\int_0^T \alpha\,\hat{x}_t(\theta, y)\, dy_t - \frac{1}{2}\int_0^T \alpha^2\,\hat{x}_t(\theta, y)^2\, dt \right). \tag{2.6}$$

The verification of this fact is quite straight forward; for example it follows from

the combination of two results (Theorems 7.3.1 and 7.3.2, pp 176) from Kallianpur (1980).

Let $\hat{\theta}_T(y)$ be the m.l.e. of θ based on observations (y_t), $0 \le t \le T$ i.e., the maximum, over the parameter space, of the above likelihood ratio is attained at $\hat{\theta}_T$. Suppose H_θ, R_θ and G_θ are trace class operators (R_θ is self-adjoint) defined on $L^2[0,T]$ with respective kernels :

$$H_\theta(t,s) = \frac{\alpha}{\sqrt{\alpha^2+\beta^2}}\,[\,1+\beta(t-s)-\sqrt{\alpha^2+\beta^2}(t-s)\,]\,e^{-\sqrt{\alpha^2+\beta^2}(t-s)} \tag{2.7}$$

if $0 \le s \le t \le T$ and equal to zero otherwise;

$$R_\theta(t,s) = \frac{e^{-\beta|t-s|} - e^{-\beta(t+s)}}{2\beta},\ 0 \le s,t \le T; \tag{2.8}$$

and

$$G_\theta(t,s) = (1+\frac{\beta}{\sqrt{\alpha^2+\beta^2}})\,[\,1+\beta(t-s)\,]\,e^{-\sqrt{\alpha^2+\beta^2}(t-s)} \tag{2.9}$$

if $0 \le s \le t \le T$ and equal to zero otherwise.

Let H^* and G^* denote the corresponding adjoint operators. Then it is easy to verify the following:

$$i)\ \lim_{T\to\infty}\frac{1}{T}\,trace\,[HRH^* + HH^*] = \sigma_1^2 < \infty$$

$$ii)\ \lim_{T\to\infty}\frac{1}{T}\,trace\,[GRG^* + GG^*] = \sigma_2^2 < \infty$$

and

$$iii)\ \lim_{T\to\infty}\frac{1}{T}\,trace\,[HRG^* + HG^*] = \sigma_{12} < \infty .$$

Also if Σ denotes the 2×2 symmetric matrix with $\Sigma_{11} = \sigma_1^2, \Sigma_{22} = \sigma_2^2$, and $\Sigma_{12} = \sigma_{12}$, then it is easy to check that Σ is strictly positive definite.

Theorem 2.1 *The m.l.e. of θ has the following properties:*

i) $\hat{\theta}_T$ is a strongly cosistent estimator of θ.

ii) As T tends to infinity the distribution of $\sqrt{T}\,(\hat{\theta}_T-\theta)$ converges to the normal distribution with zero mean and covariance matrix Σ^{-1}. Furthemore, for every $p>0$, the p th moment of the norm of $\sqrt{T}\,(\hat{\theta}_T-\theta)$ converges to the pth moment of the norm of this normal variable.

iii) For $h>0$ and large T, $T>T_0$,

$$P_\theta^T\{\,|\sqrt{T}\,(\hat{\theta}_T-\theta)|> h\,\} = B_0\exp(-b_0h^2)$$

where $B_0, b_0>0$ are constants.

3. Proof: The result is proved by verifying the conditions of the theorem given below (Theorem 3.1) which is a modified version of a result by Ibragimov & Hasminski(1981) (Theorem 10.1, Ch.1). In this version the conditions of the theorem are stated in terms of the loglikelihood function rather than the likelihood function and moreover the statement is simplified to suit this particular example of bivariate parameter estimation. The proof of this modified version can be easily deduced from a more general result from Kallianpur and Selukar(1989).

For stating Theorem 3.1 we first need some notation:

For $\theta \in \Theta$ and observations (y_t), $0 \le t \le T$; let

$$l(\theta) = l_T(\theta,y) = \ln\frac{dP_\theta^T}{d\mu_W}(y). \tag{3.1}$$

Then from (2.6),

$$l(\theta) = \int_0^T \alpha\,\hat{x}_t(\theta,y)\,dy_t - \frac{1}{2}\int_0^T \alpha^2\,\hat{x}_t(\theta,y)^2\,dt. \tag{3.2}$$

Suppose that $\theta=(\alpha,\beta)$, an element of Θ, is the true parameter. For each $T>0$, define a random function $Z_T(u)$ with domain U_T, a subset of R^2, as follows:

$$U_T = \sqrt{T}\,(\Theta - \theta) = \sqrt{T}\,(a-\alpha, b-\alpha) \times (a-\beta, b-\beta); \tag{3.3}$$

and for $u \in U_T$,

$$Z_T(u) = l(\theta + \frac{u}{\sqrt{T}}) - l(\theta) \tag{3.4}$$

(Clearly Z_T and U_T depend on the true parameter θ but this dependence is suppressed for notational convenience.)

Theorem 3.1 *Assume that the random functions $Z_T(u)$ satisfy the following three conditions:*

I)

$$\sup_{\substack{|u|, |w| \leq M \\ u, w \in U_T}} E_\theta |Z_T(u) - Z_T(w)|^4 \leq C M^4 |u - w|^4$$

II) For $u \in U_T$ and T large, $T > T_0$,

$$E_\theta \exp(\frac{1}{4} Z_T(u)) \leq \exp(-C' |u|^2).$$

III) As $T \to \infty$ the finite dimensional distributions of $Z_T(u)$ converge to the finite dimensional distributions of $Z(u)$ where for $u \in R^2$,

$$Z(u) = u' Y_\theta - \frac{1}{2} u' \Delta_\theta u;$$

here Y_θ is a zero mean bivariate normal variable with invertible covariance matrix Δ_θ. (Note that $Z(u)$ is a real valued, continuous random function defined on R^2 which attains its maximum at a unique (random) point $\Delta^{-1} Y_\theta$.)

Then the m.l.e. $\hat{\theta}_T$ is a consistent estimator of θ and $\sqrt{T}\,(\hat{\theta}_T - \theta)$ converges, in distribution, to $\Delta_\theta^{-1} Y_\theta$. Moreover, for $h > 0$ and large T, $T > T_0$,

$$P_\theta^T \{ |\sqrt{T}\,(\hat{\theta}_T - \theta)| > h \} = B_0 \exp(-b_0 h^2)$$

where $B_0, b_0 > 0$ are constants. This implies that for all $p \geq 0$,

$$\lim_{T \to \infty} E |\sqrt{T}\,(\hat{\theta}_T - \theta)|^p = E |\Delta^{-1} Y_\theta|^p.$$

We shall verify the conditions of Theorem 3.1 using several lemmas. The first two lemmas are technical; Lemma 3.1, (i) gives a bound on the higher moments of the L^2 norm of a square integrable Gaussian process in terms of its second moment and Lemma 3.1, (ii) bounds its moment generating function. The lemma is proved using some simple properties of the Karhunen-Loeve expansion of the square integrable Gaussian processes. The details of the proof can be found in Selukar(1989) (Lemmas 4.1 & 4.2, Ch.3).

Let $(Y_t), 0 \leq t \leq T$ be a *zero* mean Gaussian process such that

$$E \int_0^T Y_t^2 \, dt < \infty .$$

Let $D(t,s)$ be the covariance function of the process and D be the corresponding covarince operator. It is well-known that D is a self-adjoint trace class operator and

$$trace\,(D) = E \int_0^T (Y_t)^2 \, dt. \tag{3.5}$$

Lemma 3.1 *i) For all* $k \geq 1$

$$E\left(\int_0^T Y_t^2 \, dt\right)^k < k^k \left[E\left(\int_0^T Y_t^2 \, dt\right)\right]^k .$$

ii)

$$E \exp\left(-\int_0^T Y_t^2 \, dt\right) \leq \exp\left(-\frac{trace\,(D)}{1+2\|D\|}\right)$$

where $\|D\|$ *is the operator norm of* D.

In the next lemma we collect some useful properties of an integral operator defined on $L^2[0,T]$. The proof of this lemma is simple and so it is ommitted.

For $\lambda > 0$ and m a non-negative integer, let L be an integral operator defined on $L^2[0,T]$ with kernel $L(t,s)$ given by,

$$L(t,s) = (t-s)^m e^{-\lambda(t-s)}, \quad 0 \le s \le t \le T$$

$$= 0 \text{ otherwise.}$$

That is, for $f \in L^2[0,T]$,

$$(Lf)(t) = \int_0^T L(t,s) f(s)\, ds.$$

L is a special case of *Volterra* operator. Let $\|L\|$ be the operator norm of L and L^* denote the adjoint of L. Then LL^* is a self-adjoint trace class operator and it is easy to check that

$$Trace(LL^*) = \int_0^T \int_0^t L^2(t,s)\, ds\, dt$$

$$= T \int_0^T u^{2m} e^{-2\lambda u}\, du - \int_0^T u^{2m+1} e^{-2\lambda u}\, du.$$

Lemma 3.2

i) $\|L\| \le \sqrt{2}\left(\frac{\Gamma(m+1)}{\lambda^{m+1}}\right)$

ii) $Trace(LL^*) \le \frac{\Gamma(2m+1)}{(2\lambda)^{2m+1}} T$

iii) For large T, $Trace(LL^*) \ge \frac{\Gamma(2m+1)}{(2\lambda)^{2m+1}} \frac{T}{2}$

iv) $\lim_{T \to \infty} \frac{1}{T} Trace(LL^*) = \frac{\Gamma(2m+1)}{(2\lambda)^{2m+1}}$

v) If L_1 *and* L_2 *are two Volterra operators of the above type then the operator* $L_1 L_2$ *is also a Volterra operator which is a finite linear combination of the operators of the above type.*

Remark 3.1 Note that the operator norm of L has a bound *independent* of T and the trace of LL^* is of the same order as that of T. It is obvious that statements of the above lemma can also be obtained for an operator M which is a linear combination of L_is.

Let $\hat{X}_t(\theta,y) = \alpha\,\hat{x}_t(\theta,y)$. Then, using (2.4),

$$\hat{X}_t(\theta,y) = (-\beta + \sqrt{\alpha^2+\beta^2}) \int_0^t e^{-\sqrt{\alpha^2+\beta^2}(t-s)}\, dy_s . \tag{3.6}$$

From (3.2)

$$l(\theta,y) = \int_0^T \hat{X}_t(\theta,y)\, dy_t - \frac{1}{2}\int_0^T \hat{X}_t^{\,2}(\theta\, y)\, dt$$

Therefore

$$\begin{aligned} Z_T(u) &= l(\theta+\frac{u}{\sqrt{T}}) - l(\theta) \\ &= \int_0^T [\hat{X}_t(\theta+\frac{u}{\sqrt{T}},y)-\hat{X}_t(\theta,y)]\, dy_t \\ &\quad - \frac{1}{2}\int_0^T [\hat{X}_t^{\,2}(\theta+\frac{u}{\sqrt{T}},y)-\hat{X}_t^{\,2}(\theta,y)]\, dt. \end{aligned}$$

If we complete the square in the second term of the RHS and rearrange the terms we get,

$$\begin{aligned} Z_T(u) &= \int_0^T [\hat{X}_t(\theta+\frac{u}{\sqrt{T}},y) - \hat{X}_t(\theta,y)]\, d\nu_t^{\theta} \\ &\quad - \frac{1}{2}\int_0^T [\hat{X}_t(\theta+\frac{u}{\sqrt{T}},y) - \hat{X}_t(\theta,y)]^2\, dt \end{aligned} \tag{3.7}$$

where $d\nu_t^{\theta} = dy_t - \hat{X}_t(\theta,y)\, dt$. Recall that under P_θ^T, ν_t^{θ} is a standard Wiener process.

From now on, unless stated otherwise, all the expectations are taken w.r.t. the true probability measure P_θ^T. Also, in order to simplify the notation we may sometimes write, $Z(u) = Z_T(u)$, $\hat{X}_t(u) = \hat{X}_t(\theta+\frac{u}{\sqrt{T}},y)$ and $d\nu_t = d\nu_t^{\theta}$. The next lemma verifies the first condition of Theorem 3.1.

Lemma 3.3 *For* $u,w \in U_T$, $|u|, |w| \le M$,

$$E\ (Z_T(u) - Z_T(w)\ )^4 \le C\ M^4\ |u - w|^4$$

where C *depends only on* a *and* b. *Recall that* $\Theta = (a,b) \times (a,b)$.

Proof From (3.7),

$$Z_T(u) - Z_T(w) = \int_0^T [\ \hat{X}_t(u) - \hat{X}_t(w)\ ]\ d\nu_t$$

$$- \frac{1}{2}\int_0^T \{\ [\ \hat{X}_t(u) - \hat{X}_t(0)\ ]^2 - [\ \hat{X}_t(w) - \hat{X}_t(0)\ ]^2\ \}\ dt$$

$$= TERM\ 1 - \frac{1}{2} TERM\ 2 \quad \text{say.}$$

Then,

$$E(Z(u) - Z(w))^4 \le 16\{\ E(TERM\ 1)^4 + \frac{1}{16} E(TERM\ 2)^4\ \}. \tag{3.8}$$

Consider $E(TERM\ 1)^4$:

$$E(TERM\ 1)^4 = E\{\int_0^T [\ \hat{X}_t(u) - \hat{X}_t(w)\ ]\ d\nu_t\ \}^4$$

$$\le 16E\ \{\int_0^T [\ \hat{X}_t(u) - \hat{X}_t(w)\ ]^2\ dt\ \}^2$$

(Follows from Burkholder's martingale inequality and the fact that $(\ \nu_t\ )$ is a Wiener process under P_θ^T.)

$$\le 64\{\int_0^T E\ [\ \hat{X}_t(u) - \hat{X}_t(w)\ ]^2\ dt\ \}^2 \tag{3.9}$$

because of Lemma 3.1 (i).

Note that from (3.6),

$$\hat{X}_t(u) - \hat{X}_t(w) = \int_0^t (L_u - L_w)(t,s)\ dy_s \tag{3.10}$$

where $L_u = L_{\theta + \frac{u}{\sqrt{T}}}$ is given as follows:

$$L_u(t,s) = A_u\, e^{-B_u(t-s)} \quad \text{for } 0 \le s \le t \le T$$

$$= 0 \text{ otherwise;}$$

$$A_u = A_{\theta + \frac{u}{\sqrt{T}}} = -(\beta + \frac{u_2}{\sqrt{T}}) + \{(\beta + \frac{u_2}{\sqrt{T}})^2 + (\alpha + \frac{u_1}{\sqrt{T}})^2\}^{\frac{1}{2}}$$

and

$$B_u = B_{\theta + \frac{u}{\sqrt{T}}} = \{(\beta + \frac{u_2}{\sqrt{T}})^2 + (\alpha + \frac{u_1}{\sqrt{T}})^2\}^{\frac{1}{2}}.$$

(Note that $u = (u_1, u_2)$ is a point in $U_T \subset R^2$.)

Now recall that (see (1.2))

$$dy_t = \alpha x_t\, dt + dW_2(t) \tag{3.11}$$

where the 'signal' (x_t) and the observation 'noise' $(W_2(t))$ are independent. Let $R(t,s)$ denote the covariance function of x_t i.e. $R(t,s) = E(x_t x_s)$. Then since x_t is the familiar Ornstein-Uhlenbeck process it follows that

$$R(t,s) = \frac{e^{-\beta|t-s|} - e^{-\beta(t+s)}}{2\beta}. \tag{3.12}$$

If R is the integral operator with the kernel $R(t,s)$. Then it is easy to check that $R = R_1 R_1^*$ where R_1 is the Volterra operator with kernel

$$R_1(t,s) = e^{(-\beta(t-s))} \quad \text{for } 0 \le s \le t \le T.$$

Therefore, by Lemma 3.2, R is a trace class operator and

$$\|R\| \le \frac{2}{\beta^2}. \tag{3.13}$$

From (3.10), (3.11) and the independence of (x_t) and $(W_2(t))$ it follows that

$$\int_0^T E[\hat{X}_t(u) - \hat{X}_t(w)]^2\, dt$$

$$= \int_0^T \int_0^t \int_0^t (L_u - L_w)(t,v)\, R(v,s) (L_u - L_w)(t,s)\, ds\, dv\, dt$$

$$+ \int_0^T \int_0^t (L_u - L_w)^2(t,s)\, ds\, dt.$$

$$= Trace\,\{(L_u - L_w) R (L_u - L_w)^* \} + Trace\,\{(L_u - L_w)(L_u - L_w)^* \}$$

where L_u denotes the integral operator with kernel $L_u(t,s)$ and $(L_u)^*$ its adjoint.

Next using (3.13) and the fact that, for any two trace class operators J_1 and J_2,

$$Trace(J_1 J_2) \le Trace(J_1).\, ||J_2|| \tag{3.14}$$

where $||J_2|| =$ the operator norm of J_2, it follows that

$$\int_0^T E\,[\hat{X}_t(u) - \hat{X}_t(w)]^2\, dt \le (1 + \frac{2}{\beta^2})\, Trace(L_u - L_w)(L_u - L_w)^*$$

$$= (1 + \frac{2}{\beta^2}) \int_0^T \int_0^t (L_u - L_w)^2(t,s)\, ds\, dt.$$

$$\le C\, \{(u_1 - w_1)^2 + (u_2 - w_2)^2\} \tag{3.15}$$

The last step is obtained using Taylor's theorem and Lemma 3.2 as follows:

$$L_u(t,s) - L_w(t,s) = \frac{(u_1 - w_1)}{\sqrt{T}} \frac{\partial}{\partial \alpha} L_{\underline{u}}(t,s)$$

$$+ \frac{(u_2 - w_2)}{\sqrt{T}} \frac{\partial}{\partial \beta} L_{\underline{u}}(t,s)$$

where $\underline{u} \in (\alpha + \frac{u_1}{\sqrt{T}}, \alpha + \frac{w_1}{\sqrt{T}}) \times (\beta + \frac{u_2}{\sqrt{T}}, \beta + \frac{w_2}{\sqrt{T}}) \subset \Theta$. ($\underline{u}$ may depend on t and s.) Therefore

$$(L_u - L_w)^2(t,s)$$

$$\le \frac{4}{T} [(u_1 - w_1)^2 + (u_2 - w_2)^2] \sup_{\theta \in \Theta} \{(\frac{\partial}{\partial \alpha} L_\theta)^2 + (\frac{\partial}{\partial \beta} L_\theta)^2(t,s)\}$$

$$\leq \frac{4}{T}[(u_1-w_1)^2+(u_2-w_2)^2]\{(C_1+C_2(t-s)+C_3(t-s)^2)\exp(-C_4(t-s))\}$$

C_is are positive constants which depend on $a>0$ and b only. (Recall that $\Theta=(a,b)\times(a,b)$) Therefore

$$(1+\frac{2}{\beta^2})\int_0^T\int_0^t (L_u-L_w)^2(t,s)\,ds\,dt$$

$$\leq \frac{4}{T}[(u_1-w_1)^2+(u_2-w_2)^2]$$

$$\times\int_0^T\int_0^t \{(C_1+C_2(t-s)+C_3(t-s)^2)\exp(-C_4(t-s))\}\,ds\,dt$$

$$\leq C\,\frac{||u-w||^2}{T}\times T$$

(Follows from Lemma 3.2 (ii) and Remark 3.1) Thus, from (3.9) and (3.15) it follows that

$$E\,(TERM\ 1)^4 \leq C\ ||u-w||^4. \tag{3.16}$$

Now consider $TERM\ 2$:

$$(TERM\ 2)^4 = \int_0^T \{\,[\hat{X}_t(u)-\hat{X}_t(0)]^2-[\hat{X}_t(w)-\hat{X}_t(0)]^2\,\}\,dt$$

$$\leq (\int_0^T ([\hat{X}_t(u)-\hat{X}_t(0)]+[\hat{X}_t(w)-\hat{X}_t(0)])^2\,dt)^2$$

$$\times(\int_0^T (\hat{X}_t(u)-\hat{X}_t(w))^2\,dt)^2.$$

We have used the identity $(x^2-y^2)=(x+y)(x-y)$ and then applied the Cauchy-Schwartz inequality. Thus, again applying the Cauchy-Schwartz inequality,

$$E\,(TERM\ 2)^4 \leq \{E\,(\int_0^T ([\hat{X}_t(u)-\hat{X}_t(0)]+[\hat{X}_t(w)-\hat{X}_t(0)])^2\,dt)^4$$

$$\times E\,(\int_0^T (\hat{X}_t(u)-\hat{X}_t(w))^2\,dt)^4\}^{\frac{1}{2}}.$$

$$\leq \{16(C4^4[||u-0||^8+||w-0||^8]) \times C\ ||u-w||^8\}^{\frac{1}{2}}.$$

(Using Lemma 3.1 (i) and the calculations done for *TERM* 1.) Thus, since $||u||, ||w|| < M$,

$$E(TERM\ 2)^4 < C\ M^4\ ||u-w||^4. \tag{3.17}$$

Now the statement of the lemma follows from (3.8), (3.16) and (3.17).

Lemma 3.4 For $u \in U_T$ *and* T *large,*

$$E\exp(\frac{1}{4}Z_T(u)) \leq \exp(-C'||u||^2)$$

where $C' > 0$, *depends on* a *and* b *only.*

Proof : From (3.7)

$$E\exp(\frac{1}{4}Z_T(u))$$

$$= E\exp(\frac{1}{4}\int_0^T [\hat{X}_t(u)-\hat{X}_t(0)]\,d\nu_t^{\theta} - \frac{1}{8}\int_0^T [\hat{X}_t(u)-\hat{X}_t(0)]^2 dt)$$

$$\leq \{E\exp(\frac{1}{2}\int_0^T [\hat{X}_t(u)-\hat{X}_t(0)]\,d\nu_t^{\theta} - \frac{1}{8}\int_0^T [\hat{X}_t(u)-\hat{X}_t(0)]^2 dt)$$

$$\times E\exp(-\frac{1}{8}\int_0^T [\hat{X}_t(u)-\hat{X}_t(0)]^2 dt)\}^{\frac{1}{2}}$$

(By the application of the Cauchy-Schwartz inequality.) Note that the first term of the product in the bracket is a density (w.r.t. P_θ^T). Therefore

$$E\exp(\frac{1}{4}Z_T(u)) \leq \{1 . E\exp(-\frac{1}{8}\int_0^T [\hat{X}_t(u)-\hat{X}_t(0)]^2 dt)\}^{\frac{1}{2}}.$$

The RHS above can be easily bounded by applying Lemma 3.1 (ii) since $\xi_t = [\hat{X}_t(u)-\hat{X}_t(0)]$ is a Gaussian process. Let $F(t,s)$ be the covariance

function of this process and F be the corresponding covariance operator. Then

$$E\exp\left(\frac{1}{4}Z_T(u)\right) \le \exp\left(-\frac{1}{16}\frac{traceF}{1+2\|F\|}\right).$$

From the calculations made in order to bound $E\,(TERM\ 1)^2$ in Lemma 3.3 (with $u = u$ and $w = 0$) we can make the following obsevations:

i) $F = (L_u - L_0)R\,(L_u - L_0)^* + (L_u - L_0)(L_u - L_0)^*$

ii) $\|F\| \le C'$

iii) $TrF = Tr(L_u - L_0)R\,(L_u - L_0)^* + Tr(L_u - L_0)(L_u - L_0)^*$ and since the first term is always positive,

$$TraceF \ge Trace(L_u - L_0)(L_u - L_0)^* \ge C\,\|u\|^2;$$

where $C > 0$.

In order to see (iii) first recall that

$$Trace(L_u - L_0)(L_u - L_0)^* = \int_0^T\int_0^t (L_u - L_0)^2(t,s)\,ds\,dt.$$

Now

$$L_u(t,s) - L_0(t,s) = \frac{(u_1)}{\sqrt{T}}\frac{\partial}{\partial\alpha}L_{\underline{u}}(t,s)$$

$$+\ \frac{(u_2)}{\sqrt{T}}\frac{\partial}{\partial\beta}L_{\underline{u}}(t,s)$$

where $\underline{u} \in (\alpha+\frac{u_1}{\sqrt{T}},\alpha)\times(\beta+\frac{u_2}{\sqrt{T}},\beta) \subset \Theta$. ($\underline{u}$ may depend on t and s.) Therefore

$$(L_u - L_0)^2(t,s)$$

$$\ge \frac{4}{T}[(u_1)^2 + (u_2)^2]\left\{\inf_{\theta\in\Theta}\left(\frac{\partial}{\partial\alpha}L_\theta\right)^2 + \inf_{\theta\in\Theta}\left(\frac{\partial}{\partial\beta}L_\theta\right)^2(t,s)\right\}$$

$$\ge \frac{4}{T}[(u_1)^2 + (u_2)^2]\{(C_1 + C_2(t-s) + C_3(t-s)^2)\exp(-C_4(t-s))\};$$

C_is are positive constants which depend on $a > 0$ and b only. (Recall that $\Theta = (a,b) \times (a,b)$) Therefore

$$\int_0^T \int_0^t (L_u - L_0)^2(t,s)\, ds\, dt$$

$$\geq \frac{4}{T} \|u\|^2 \int_0^T \int_0^t \{(C_1 + C_2(t-s) + C_3(t-s)^2) \exp(-C_4(t-s))\}\, ds\, dt$$

$$\geq \frac{4}{T} \|u\|^2 C \,.\, T = C \,.\, \|u\|^2.$$

for large T. (Follows from Lemma 3.2 (iii) and Remark 3.1) Thus (iii) follows and we finally get that, for large T,

$$E \exp\left(\frac{1}{4} Z_T(u)\right) \leq \exp(-C' \|u\|^2)$$

where $C' > 0$, depends on a and b only. This is the statement of the lemma.

For $u \in R^2$ let

$$Z_\theta(u) = u' Y_\theta - \frac{1}{2} u' \Sigma u \tag{3.18}$$

where Y_θ is a zero mean bivariate Normal with covariance matrix Σ.

Now we will show that finite dimensional distributions of $Z_T(u)$ converge to the finite dimensional distributions of $Z_\theta(u)$ as $T \to \infty$. This together with Lemmas 3.3 and 3.4 verify all the conditions of Theorem 3.1 and hence prove Theorem 2.1.

The convergence of finite dimensional distributions is shown in two steps. First we define random functions $Z^*_T(u)$ (with domains U_T) such that for any fixed $u \in U_T$,

$$E(Z_T(u) - Z^*_T(u))^2 \to 0 \text{ as } T \to \infty. \tag{3.19}$$

Next we show that finite dimensional distributions of $Z^*_T(u)$ converge to the

finite dimensional distributions of $Z_\theta(u)$.

These two steps are clearly sufficient for our purpose.

Let us begin the first step :

For $u \in U_T$ let

$$Z^*_T(u) = \int_0^T [\frac{u_1}{\sqrt{T}} h_t(\theta) + \frac{u_2}{\sqrt{T}} g_t(\theta)]\, d\nu_t^\theta$$

$$- \frac{1}{2}\int_0^T [\frac{u_1}{\sqrt{T}} h_t(\theta) + \frac{u_2}{\sqrt{T}} g_t(\theta)]^2\, dt. \tag{3.20}$$

where

$$h_t(\theta) = \int_0^t \frac{\partial}{\partial\alpha} L_\theta(t,s)\, dy_s$$

$$= \int_0^t \frac{\alpha}{\sqrt{\alpha^2+\beta^2}} [1+\beta(t-s) - \sqrt{\alpha^2+\beta^2}(t-s)]\, e^{-\sqrt{\alpha^2+\beta^2}(t-s)}\, dy_s$$

$$= \int_0^t H_\theta(t,s)\, dy_s \tag{3.21}$$

and

$$g_t(\theta) = \int_0^t \frac{\partial}{\partial\beta} L_\theta(t,s)\, dy_s$$

$$= \int_0^t (1+\frac{\beta}{\sqrt{\alpha^2+\beta^2}})\, [1+\beta(t-s)]\, e^{-\sqrt{\alpha^2+\beta^2}(t-s)}\, dy_s.$$

$$= \int_0^t G_\theta(t,s)\, dy_s. \tag{3.22}$$

(Recall the two integral operators H_θ & G_θ with kernels $H_\theta(t,s)$ and $G_\theta(t,s)$ defined in Section 2 (see (2.7) and (2.9)) it turns out that $H_\theta(t,s) = \frac{\partial}{\partial\alpha} L_\theta(t,s)$ and $G_\theta(t,s) = \frac{\partial}{\partial\beta} L_\theta(t,s)$.)

Lemma 3.5 For $u \in U_T$,

$$E\,(Z_T(u) - Z^*_T(u))^2 \rightarrow 0 \quad \text{as} \quad T \quad \infty.$$

Proof This proof is almost identical to that of Lemma 3.3. From (3.7) and (3.21)

$$Z_T(u) - Z_T^*(u) = TERMA - \frac{1}{2} TERMB \tag{3.23}$$

where

$$TERMA = \int_0^T [\hat{X}_t(u) - \hat{X}_t(0) - (\frac{u_1}{\sqrt{T}} h_t(\theta) + \frac{u_2}{\sqrt{T}} g_t(\theta))]\, d\nu_t^{\theta}$$

and

$$TERMB = \int_0^T ([\hat{X}_t(u) - \hat{X}_t(0)]^2 - [\frac{u_1}{\sqrt{T}} h_t(\theta) + \frac{u_2}{\sqrt{T}} g_t(\theta)]^2)\, dt.$$

Let us consider $E(TERMA)^2$. Since (ν_t) is a Wiener process,

$$E(TERMA)^2 = E\int_0^T [\hat{X}_t(u) - \hat{X}_t(0) - (\frac{u_1}{\sqrt{T}} h_t(\theta) + \frac{u_2}{\sqrt{T}} g_t(\theta))]^2\, dt$$

$$= E\int_0^T [\int_0^t (L_u - L_0 - (\frac{u_1}{\sqrt{T}} H + \frac{u_2}{\sqrt{T}} G))(t,s)\, dy_s\,]^2\, dt$$

$$\leq (1 + \frac{2}{\beta^2}) \int_0^T \int_0^t (L_u - L_0 - (\frac{u_1}{\sqrt{T}} H + \frac{u_2}{\sqrt{T}} G))^2(t,s)\, ds\, dt$$

The last step is obtained using the same arguments as in (3.7) - (3.15). Now note that, using Taylor expansion,

$$(L_u - L_0 - (\frac{u_1}{\sqrt{T}} H + \frac{u_2}{\sqrt{T}} G))(t,s) = (\frac{u_1^2}{2T} \frac{\partial^2}{\partial \alpha^2} L_{\underline{u}} + \frac{u_2^2}{2T} \frac{\partial^2}{\partial \beta^2} L_{\underline{u}}$$

$$+ \frac{u_1 u_2}{2T} \frac{\partial^2}{\partial \alpha \partial \beta} L_{\underline{u}})(t,s)$$

for some $\underline{u} \in (\theta, \theta + \frac{u}{\sqrt{T}}) \subset \Theta$. Therefore

$$(L_u - L_0 - (\frac{u_1}{\sqrt{T}} H + \frac{u_2}{\sqrt{T}} G))^2(t,s)$$

$$\leq \frac{C}{T^2}(u_1{}^2 + u_2{}^2)^2 \sup_{\theta \in \Theta} [(\frac{\partial^2}{\partial \alpha^2} L_\theta)^2 + (\frac{\partial^2}{\partial \beta^2} L_\theta)^2 + (\frac{\partial^2}{\partial \alpha \partial \beta} L_\theta)^2](t,s)$$

$$\leq C \frac{||u||^4}{T^2} [\sum_{i=0}^{4} (t-s)^i] \exp(-C_1(t-s)).$$

(The positive constants C and C_1 depend on a and b only.) Therefore

$$(1 + \frac{2}{\beta^2}) \int_0^T \int_0^t (L_u - L_0 - (\frac{u_1}{\sqrt{T}} H + \frac{u_2}{\sqrt{T}} G))^2(t,s)\, ds\, dt$$

$$\leq C \frac{||u||^4}{T^2} \int_0^T \int_0^t [\sum_{i=0}^{4} (t-s)^i] \exp(-C_1(t-s))\, ds\, dt$$

$$\leq C \frac{||u||^4}{T^2} . T \to 0 \text{ as } T \to \infty .$$

(since $u \in U_T$ is fixed) Thus $E(TERMA)^2 \to 0$ as $T \to \infty$.

The steps to show that $E(TERMB)^2 \to 0$ as $T \to \infty$ are also very similar. Using the Cauchy - Schwartz inequality we get

$$(TERMB)^2 \leq (\int_0^T [\hat{X}_t(u) - \hat{X}_t(0) + (\frac{u_1}{\sqrt{T}} h_t(\theta) + \frac{u_2}{\sqrt{T}} g_t(\theta))]^2 dt)$$

$$\times (\int_0^T [\hat{X}_t(u) - \hat{X}_t(0) - (\frac{u_1}{\sqrt{T}} h_t(\theta) + \frac{u_2}{\sqrt{T}} g_t(\theta))]^2 dt)$$

Again applying the Cauchy - Schwartz inequality we get

$$E(TERMB)^2 \leq \{E(\int_0^T [\hat{X}_t(u) - \hat{X}_t(0) + (\frac{u_1}{\sqrt{T}} h_t(\theta) + \frac{u_2}{\sqrt{T}} g_t(\theta))]^2 dt)^2$$

$$\times E(\int_0^T [\hat{X}_t(u) - \hat{X}_t(0) - (\frac{u_1}{\sqrt{T}} h_t(\theta) + \frac{u_2}{\sqrt{T}} g_t(\theta))]^2 dt)^2\}^{\frac{1}{2}}$$

Since all the processes involved are Gaussian we can use Lemma 3.1 (i) and then with some simple manipulations we get

$$E\,(TERMB\,)^2$$

$$\leq\ C\ \{(E\int_0^T [\hat{X}_t(u)-\hat{X}_t(0)]^2\,dt)^2+\frac{\|u\|^4}{T^2}(E\int_0^T [h^2_t(\theta)+g^2_t(\theta)]\,dt)^2\}^{\frac{1}{2}}$$

$$\times\, E\,(TERMA\,)^2.$$

Taking $w = 0$ in the calculations of $E\,(TERM\ 1)^4$ in Lemma 3.3 we get

$$(E\int_0^T [\hat{X}_t(u)-\hat{X}_t(0)]^2\,dt)^2 \leq\ C\ \|u\|^4.$$

Also, it is easy to verify that

$$(E\int_0^T [h^2_t(\theta)+g^2_t(\theta))]\,dt)^2 \leq\ C\ T^2.$$

Therefore we get that

$$E\,(TERMB\,)^2 \leq\ C\ \|u\|^2 E\,(TERMA\,)^2\ \rightarrow\ 0\ \text{ as }\ T\ \rightarrow\ \infty.$$

This concludes the proof of Lemma 3.5.

Now let us begin the second step :

First we state a version of Central limit theorem which is useful for our purpose (see, Basawa & Prakasa Rao (1980), Theorem 2.1, Appendix 2, pp 405).

Let $\{W(t), t\geq 0\}$ denote the standard m dimensional Brownian motion. Suppose that $F(s) = ((f_{kj}(s)))_{n\times m}$ is a random matrix valued function such that its elements $f_{kj} \in H[0,T]$ for all $T > 0$. (A random function $f \in H[0,T]$ iff it is adapted to the Wiener filtration and

$$E\int_0^T f^2(t)\,dt < \infty .)$$

Set $\mathbf{f}_k(s) = (f_{k1}(s), \ldots, f_{km}(s))$, $1 \le k \le n$.

Theorem B & P : *Suppose that the random matrix valued function $F(s)$ satisfies the following condition :*

$$\frac{1}{T}\int_0^T < \mathbf{f}_k(s), \mathbf{f}_j(s) > ds \to c_{kj}$$

in probability as $T \to \infty$ where $c_{kj}, 1 \le k, j \le n$ are finite. Then the distribution of

$$T^{-\frac{1}{2}} \int_0^T F(s)\, dW(s)$$

converges to the Normal distribution with mean zero and covariance matrix $C = (c_{kj})$ as $T \to \infty$.

From the above result, the fact that (v_t) is a Wiener process and the special forms of $Z^*_T(u)$ and $Z_\theta(u)$ it is clear that we only have to show the following:

i) $\frac{1}{T}\int_0^T h^2_t(\theta)\, dt \to \sigma_1^2$

ii) $\frac{1}{T}\int_0^T g^2_t(\theta)\, dt \to \sigma_2^2$

iii) $\frac{1}{T}\int_0^T g_t(\theta)\, h_t(\theta)\, dt \to \sigma_{1,2}$ in *probability* as $T \to \infty$.

(Note that just as $Z_\theta(u)$ (see (3.18)) we can write

$$Z^*_T(u) = u' Q_T(\theta) - \frac{1}{2} u' \Delta_T(\theta)\, u$$

where $Q_T(\theta)$ is bivariate Normal random variable and $\Delta_T(\theta)$ is 2×2 random symmetric matrix described as below:

$$Q_T(\theta) = \left(\frac{1}{\sqrt{T}} \int_0^T h_t(\theta)\, dv_t,\ \frac{1}{\sqrt{T}} \int_0^T g_t(\theta)\, dv_t \right)'$$

and

$$\Delta_{T\,11} = \frac{1}{T}\int_0^T h^2{}_t(\theta)\,dt, \quad \Delta_{T\,22} = \frac{1}{T}\int_0^T g^2{}_t(\theta)\,dt,$$

$$\Delta_{T\,12} = \frac{1}{T}\int_0^T g_t(\theta)\,h_t(\theta)\,dt.$$

Showing (i), (ii) and (iii) amounts to showing that $\Delta_T(\theta)$ converges to Σ in probability which in turn implies that $Q_T(\theta)$ converges to Y_θ in distribution.)

Let us first show (i): We will show that

$$E\left(\frac{1}{T}\int_0^T h^2{}_t(\theta)\,dt - \sigma_1{}^2\right)^2 \to 0 \quad \text{as } T \to \infty.$$

Let $M_T = E\,\frac{1}{T}\int_0^T h^2{}_t(\theta)\,dt.$

Now

$$E\left(\frac{1}{T}\int_0^T h^2{}_t(\theta)\,dt - \sigma_1{}^2\right)^2$$

$$\leq 2E\left(\frac{1}{T}\int_0^T h^2{}_t(\theta)\,dt - M_T\right)^2 + 2(M_T - \sigma_1{}^2)^2.$$

Consider the first term on the RHS:

$$E\left(\frac{1}{T}\int_0^T h^2{}_t(\theta)\,dt - M_T\right)^2 = E\left(\frac{1}{T}\int_0^T h^2{}_t(\theta)\,dt\right)^2 - (M_T)^2$$

$$= \frac{1}{T^2}\int_0^T\int_0^T (E\,h^2{}_t\,h^2{}_s)\,ds\,dt - (M_T)^2$$

Let $J(s,t) = E\,h_t\,h_s$; then since (h_t) is a Gaussian process

$$E\,h^2{}_t\,h^2{}_s = 2J^2(s,t) + J(t,t)\,J(s,s).$$

This together with the fact that $M_T = \frac{1}{T}\int_0^T J(t,t)\,dt$ implies that

$$E\left(\frac{1}{T}\int_0^T h^2_t(\theta)\,dt - M_T\right)^2 = \frac{2}{T^2}\int_0^T\int_0^T J^2(s,t)\,ds\,dt$$

$$= \frac{2}{T^2}\,trace\,(JJ^*)$$

$$\leq \frac{2}{T^2}\,trace\,(J)\;\|J\|$$

where J is the integral operator corresponding to the symmetric kernel $J(s,t)$. Since $h_t = \int_0^t H_\theta(t,s)\,dy_s$, it is easy to check that $J = H\,R\,H^* + H\,H^*$. Therefore, using Lemma 3.3 (i) and (ii) we get that $\|J\| < C$ and $trace J \leq C\,T$. Thus

$$E\left(\frac{1}{T}\int_0^T h^2_t(\theta)\,dt - M_T\right)^2 \leq \frac{2}{T^2}\,C\,T \;\rightarrow\; 0$$

as $T \rightarrow \infty$. Therefore (i) is proved if we show that

$$M_T \rightarrow \sigma_1^{\,2} \text{ as } T \rightarrow \infty.$$

However note that

$$M_T = E\,\frac{1}{T}\int_0^T h^2_t(\theta)\,dt = \frac{1}{T}\int_0^T J(t,t)\,dt$$

$$= \frac{1}{T}\,trace J = \frac{1}{T}\,[\,trace\,(HRH^*) + trace\,(HH^*)\,].$$

Therefore

$$\lim_{T\rightarrow\infty} M_T = \sigma_1^{\,2}.$$

Thus (i) is proved. The verification of (ii) is exactly identical. Then only (iii) remains to be verified. Let

$$V_T = E\,\frac{1}{T}\int_0^T g_t(\theta)\,h_t(\theta)\,dt.$$

Consider

$$E\left(\frac{1}{T}\int_0^T g_t(\theta)\, h_t(\theta)\, dt - V_T\right)^2$$

$$= E\left(\frac{1}{T}\int_0^T g_t(\theta)\, h_t(\theta)\, dt\right)^2 - (V_T)^2$$

$$= \frac{1}{T^2}\int_0^T\int_0^T (Eg_t g_s h_t h_s)\, dsdt - (V_T)^2.$$

Note that if $\xi_1, \xi_2, \xi_3, \xi_4$ are jointly Normal then

$$E(\xi_1\xi_2\xi_3\xi_4) = E(\xi_1\xi_2)E(\xi_3\xi_4) + E(\xi_1\xi_3)E(\xi_2\xi_4) + E(\xi_1\xi_4)E(\xi_2\xi_3).$$

Therefore, if $I(s,t) = E(g_t g_s)$ and $K(s,t) = E(h_s g_t)$,

$$Eg_t g_s h_t h_s = I(t,s)J(t,s) + K(t,t)K(s,s) + K(t,s)K(s,t).$$

Therefore

$$\frac{1}{T^2}\int_0^T\int_0^T (Eg_t g_s h_t h_s)\, dsdt - (V_T)^2$$

$$= \frac{1}{T^2}\int_0^T\int_0^T (I(t,s)J(t,s) + K(t,t)K(s,s) + K(t,s)K(s,t))\, dsdt$$

$$- \frac{1}{T^2}\int_0^T\int_0^T (K(t,t)K(s,s))\, dsdt$$

$$= \frac{1}{T^2}\int_0^T\int_0^T (I(t,s)J(t,s) + K(t,s)K(s,t))\, dsdt.$$

$$\leq \frac{1}{T^2}\left[\left(\int_0^T\int_0^T I^2(t,s)dsdt\right)^{\frac{1}{2}}\left(\int_0^T\int_0^T J^2(t,s)dsdt\right)^{\frac{1}{2}}\right]$$

$$+ \frac{1}{T^2}\left[\left(\int_0^T\int_0^T K^2(t,s)dsdt\right)\right].$$

Using the facts from Lemma 3.2 and Remark 3.1 it is easy to see that the above terms tend to zero as $T \to \infty$. Let K be the integral operator with kernel $K(s,t)$

then

$$K = HRG^* + HG^*.$$

Thus (iii) is proved since

$$\lim_{T\to\infty} V_T = \frac{1}{T}\lim_{T\to\infty}\int_0^T K(t,t)\,dt$$

$$= \frac{1}{T}\lim_{T\to\infty} trace[\,HRG^* + HG^*\,] = \sigma_{1,2}.$$

This concludes the proof of Theorem 2.1.

Remark 3.2 Here we briefly explain the reasons why Kutoyants's method for studying the asymptotic properties of the m.l.e. of α in the model given in (1.2) can not be applied to the problem considered in this article. First let us recall the model given in (1.2):

$$dx_t = -x_t\,dt + dW_1(t), \qquad x_0 = 0$$

$$dy_t = \alpha x_t\,dt + dW_2(t), \qquad y_0 = 0; \quad \beta \in (a,b),\ a > 0. \tag{3.24}$$

(We have set $\beta = 1$ for simplicity.)

Just as before if we denote by $\hat{x}_t(y,\alpha)$ the conditional expectation of x_t given the observations up to time t and by ν_t the innovations process i.e.,

$$d\nu_t = dy_t - \alpha\hat{x}_t\,dt;$$

then ν_t is a standard Wiener process. Furthermore it is easy to see that

$$\ln\frac{dP_{\alpha_1}^T}{dP_{\alpha_0}^T}(y) = \int_0^T [\,\alpha_1\hat{x}_t(y,\alpha_1) - \alpha_0\hat{x}_t(y,\alpha_0)\,]\,d\nu_t$$

$$-\frac{1}{2}\int_0^T [\,\alpha_1\hat{x}_t(y,\alpha_1) - \alpha_0\hat{x}_t(y,\alpha_0)\,]^2\,dt$$

$$= \int_0^T [\,\eta_t(\alpha_1) - \eta_t(\alpha_0)\,]\,d\nu_t - \frac{1}{2}\int_0^T [\,\eta_t(\alpha_1) - \eta_t(\alpha_0)\,]^2\,dt; \tag{3.25}$$

where ν_t is defined by, $d\nu_t = dy_t - \alpha_0 \hat{x}_t\, dt$ and $\eta_t(\alpha) = \alpha \hat{x}_t(\alpha)$.

Let α_0 be the true parameter and let $\hat{\alpha}$ denote the m.l.e. of α_0. Kutoyants studies the asymptotic properties of $\hat{\alpha}$ using Ibragimov & Hasminski's approach and thus he too verifies conditions similar to those of Theorem 3.1, however, for verifying these conditions he uses the following special structure of the process $\eta_t(\alpha)$:

$$d\eta'_t(\alpha) = -g(\alpha)\,\eta'_t(\alpha)\,dt + \frac{\alpha}{g(\alpha)}\,d\nu_t, \quad \eta'_t(0) = 0; \tag{3.26}$$

where $\eta'_t(\alpha) = \frac{d}{d\alpha}\eta_t(\alpha)$, $g(\alpha) = \sqrt{1+\alpha^2}$ and ν_t is a Wiener process w.r.t. $P_\alpha{}^T$. Moreover, if we denote $\eta_t(\alpha_0) - \eta_t(\alpha)$ by $\xi_t(\alpha)$ then it is shown that $\xi_t(\alpha)$ satisfies the SDE,

$$d\xi_t = -g(\alpha)\,\xi_t\,dt + [g(\alpha) - g(\alpha_0)]\,dW_t \tag{3.27}$$

where W_t is a Wiener process (w.r.t. $P_{\alpha_0}{}^T$).

The relation (3.27), the form of the likelihood function (see (3.25)) and the following result of Novikov (see Liptser & Shiryayev Vol II, Lemma 17.3, pp 203) help verify the condition corresponding to the second condition of Theorem 3.1.

Proposition 3.1 *Assume that a zero mean Gaussian process ξ_t satisfies the following SDE:*

$$d\xi_t = a\,\xi_t\,dt + dW_t, \quad \xi_0 = 0,$$

where W_t is a standard Wiener process and a is a real number. Then

$$E\exp\left(-\int_0^T \xi_t^2\,dt\right) = \exp\left(-\frac{1}{2}TC_a\right)$$

where C_a is a psitive constant.

The exact value of C_a is unimportantand and so is not given (the proof of this result is not applicable for more general Gaussian processes). In the same fashion the relation (3.26) is used to verify the other conditions of Theorem 3.1.

In the case of estimation of β no similar simplifications are possible, for example here (setting $\alpha = 1$),

$$\ln \frac{dP_{\beta_1}^T}{dP_{\beta_0}^T}(y) = \int_0^T [\hat{x}_t(y,\beta_1) - \hat{x}_t(y,\beta_0)]\, d\nu_t$$

$$- \frac{1}{2}\int_0^T [\hat{x}_t(y,\beta_1) - \hat{x}_t(y,\beta_0)]^2\, dt; \qquad (3.28)$$

and we would like a relation similar to (3.26) to hold for $\hat{x}_t'(\beta) = \frac{d}{d\beta}\hat{x}_t(\beta)$. However, from the relation

$$d\hat{x}_t = -\beta\,\hat{x}_t\, dt + A_\beta\, d\nu_t;$$

it is possible to show that

$$d\hat{x}_t' = -B_\beta'\hat{x}_t'\, dt + \hat{x}_t\, dt + A'_\beta\, d\nu_t;$$

which is much more complicated equation than (3.26). In the same way there is no relation corresponding to (3.27) for the differences $\hat{x}_t(y,\beta_1) - \hat{x}_t(y,\beta_0)$ and thus Novikov's result can not be used. As the reader may recall (see Lemma 3.4), we resorted to Lemma 3.1 (ii) in verifying the second condition of Theorem 3.1.

Bibliography

Bagchi, A., and Borkar, V., (1984), *Parameter identification in infinite dimensional linear systems.* Stochastics, vol 12, 201 - 213.

Balakrishnan, A.V., (1973), *Stochastic differential systems.* Lecture notes in Economics and Mathematical systems, Springer - Verlag, Berlin.

Balakrishnan, A.V., (1976), *Applied Functional analysis.* Springer - Verlag, New York.

Basawa, I.V. and Prakasa Rao, B.L.S., (1980), *Statistical inference for stochastic processes.* Academic Press, London.

Ibragimov, I.A. and Hasminski, R.Z., (1981), *Statistical estimation: Asymptotic theory.* Springer - Verlag, New York.

Kallianpur, G., (1980), *Stochastic filtering theory,* Springer - Verlag, New York.

Kallianpur, G., and Selukar, R. S., (1989), *Estimation of Hilbert space valued parameters by the method of sieves,* submitted for publication.

Kutoyants, Yu.A., (1984), *Parameter estimation for stochastic processes.* R & E. Research and exposition in mathematics, No. 6, Heldermann Verlag Berlin.

Liptser, R.S. and Shiryayev, A.N., (1977), *Statistics of random processes I.* Springer - Verlag, Berlin.

Liptser, R.S. and Shiryayev, A.N., (1978), *Statistics of random processes II.* Springer - Verlag, Berlin.

Selukar, R.S., (1989) *On estimation of Hilbert space valued parameters*. Ph.D. thesis, University of North Carolina at Chapel Hill.